

ABSTRACT OF THE DISCLOSURE

Labeling process unit groups a continuous black pixel area as a group by determining the sequence of black pixels from the binary image data read from the image input device, and extracts bounding rectangle information about each of the grouped continuous black pixel areas. Row extracting process unit extracts row rectangle information contained in an original image from the group bounding rectangle information extracted by the labeling process unit. Punctuation mark identification unit identifies a punctuation mark contained in the row rectangle extracted by the row extracting process unit. With the configuration, the direction of a row can be automatically determined by checking the relative position of the punctuation mark in a row based on the extracted row rectangle information and the extracted bounding rectangle information.